Factors Affecting the Assessment Quality of Students' Scientific Argumentation Competence

Zhang, JingLu; Xue, Song

Publication date:
2021

FACTORS AFFECTING THE ASSESSMENT QUALITY OF STUDENTS' SCIENTIFIC ARGUMENTATION COMPETENCY

Based on a three round qualitative study in an iterative process of developing a scientific argumentation competency (SAC) assessment instrument, this paper explores how items can be written and what factors should be considered to improve SAC assessment. By analyzing data from interviews of teachers, cognitive think-aloud interviews of students, and after-test follow-up interviews of students in each round of the study, the instrument was modified and prepared for the next round of the study. We analyzed all the interview transcripts using thematic analysis. The study found that overlapped factors emerged from teachers' and students' interview, and no new factors contributing to the instrument improvement appeared in the third round of the study. There are in total 10 factors to be considered as important in promoting SAC assessment instrument design.

Key words: Scientific literacy; Measurement; Secondary school

INTRODUCTION

Argumentation is an important practice in science by which scientists generate, justify and evaluate scientific claims (Berland & McNeill, 2010). Argumentation also plays a central role in science education because it leads students toward deep learning by engaging them in the practice of constructing and evaluating scientific arguments. Students therefore need to be explicitly taught and assessed in scientific argumentation (SA) to acquire the competency to be involved in the practice. There has, however, been little study on developing instruments to assess students' ability of engaging in SA despite the growing interest in the topic (Osborne et al., 2016). Moreover, there has been even less study discussing how to design items to assess SAC and the challenges might encounter in the process (Diana, 2019). The focus of this paper is on the qualitative research performed during the whole instrument development process to explore how teachers and students interact with the instrument and the factors that should be considered to improve the assessment.

METHOD

SAC conceptualization

This study takes SAC as the competency of using scientific evidence and language to defend one's scientific claims reasonably, meanwhile using scientific evidence to evaluate the advantages and weaknesses of others' arguments. Drawing on Toulmin (1958)'s argumentation pattern, Erduran (2004)'s analytical framework, and Kuhn (2013)'s idea of developing argument competency, SAC is decomposed into three components with hypothesized increasing cognitive demand: *Identification of SA*, *Evaluation of SA* and *Production of SA*. Each component contains several elements: *claim*, *use of evidence*, *explanation* and *rebuttal*. For each element, there are indicators referring to the aspects from which the element is assessed. For example, the indicators for "use of evidence" are *relevance, accuracy* and *sufficiency*. The detailed description of elements and indicators would be progressively evaluated and modified according to the emerging evidence of its appropriateness during the research process.

The iterative process of assessment development

Drawing upon Wilson (2004)'s four building blocks approach of developing assessment, and Newton (2017)'s approach of developing validation arguments that emphasizes on both the outcome and the design process of assessment, the instrument here includes the *construct map* that articulates the competency to be measured, a written *test,* and *scoring rubric* for each item, and its design follows the "trial in the field-feedback-reflection-revision" cycle until the instrument is valid enough for the research. The research procedure and the participants are presented in Figure 1.

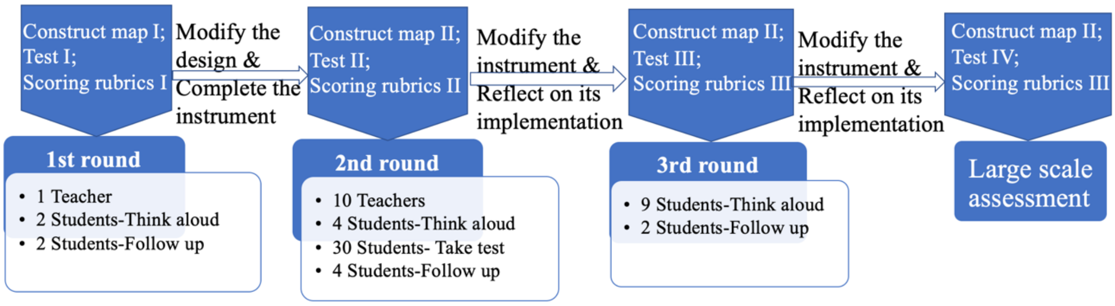


Figure 1. The research procedure

All the students were from the north of China and were in grade 11 when participating in the study. Besides, all of them were 16-17 years old. Teachers were from different places of China, two of them were science education researchers and the rest of them were Physics teachers in high school. For teachers who would like to join in a group discussion, an item panel discussion was conducted; others participated in a semi-structured interview. Teachers were given a brief research summary and the instrument around 7 days before interview, so they had enough time to read the materials. Students who participated in the think aloud interview were given the test after the interview begins. The interviewer told participants how to think aloud and provided them with an example before asking them to speak their thinking aloud. All the interview transcripts were analyzed using Nvivo. In each round of the study, teachers' and students' interviews were analyzed separately to search for emerged influencing factors, then factors were taken together for instrument modification.

RESULTS

Factors identified from the study

As shown in Figure 2, five factors emerging in the first round of the study contribute to instrument revision.

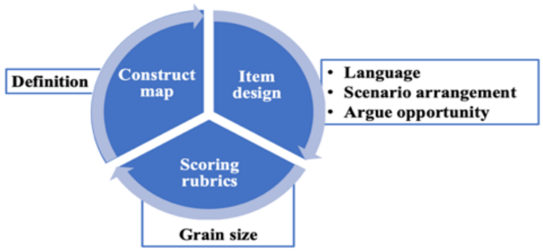


Figure 2. Factors emerged in the first round of study

Combined with factors found in the second round of the study, a total of 10 factors emerged, as shown below.

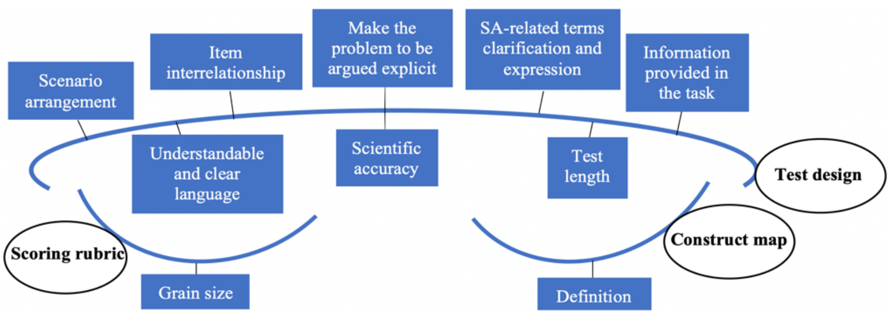


Figure 3. Factors emerged in the research

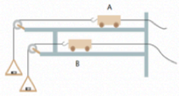
Test design before and after modification

Initially, 8 task examples, each of which contains one or more items, were designed to see how they work in the field. Each item adopting either open-ended form or matching form assesses one SAC element. The content knowledge involved in these tasks is Motion and Force that had been learned by participants. As shown in Figure 4, Task 2 assesses *the Evaluation of 'use of evidence'*.

After considering the factors that emerged from the interview data in all the three rounds of study, the modified test includes 9 tasks with three items for each SAC element. Each task, comprising of several items, focuses on one problem to be argued. An example of a task in the modified test is presented in Figure 4 Problem 8.

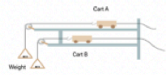
Task 2. Evaluation-use of evidence

As shown in the figure, trolley A and B start to move from standstill under the pulling force of the groove plate. They move for the same time (smooth track, smooth pulley). Student a, b and c each gives evidences to support the conclusion that "the mass of the slot code in slot B is greater than that in slot A".



a: Car A uses more environmentally friendly materials than Car B, and Car B uses thicker ropes than Car A.
b: Both cars have the same mass, and the displacement of A is greater than B.
c: Both cars passed the same displacement.
Student d thinks that the evidence given by other students cannot support the conclusion, please choose the reason in ①②③ to fill in the blanket.
a ----- reason: ( ); b ----- reason: ( ); c ----- reason: ( )
① Evidences contradict the conclusions.
② Evidences are not relevant to the conclusions.
③ Evidences are insufficient.

Problem 8: Which cart has bigger mass?



Facts about tracks and carts :
a. Cart A is solid, cart B is hollow
b. Both carts have wooden wheels
c. The track of A is wooden, and the track of B is metal
d. A is made with common material, while B is made with environmentally friendly material
e. The mass of the weight connected on the two carts are the same

The data of the two carts moving for 2 seconds under the pull of the weight:

Cart	Initial velocity V_0 (m/s)	Displacement x (m)	Movement time (s)
A	0	0.4	2
B	0	0.6	2

The friction coefficients of some surfaces:

Material	wood-ice	wood-metal	steel-ice	wood-wood
Friction coefficient	0.03	0.20	0.02	0.30

Bob: "Cart A has bigger mass. It can be judged from fact d, cart B should have used less material since it is environmentally friendly, so it is lighter."
Jane: "Environmentally friendly material has nothing to do with material's volume and mass, and environmentally friendly materials are not necessarily of smaller mass. It is true that car A has bigger mass, but it should be judged by fact a. The solid cart is definitely heavier than the hollow cart."

1. What is Bob's reason? ( )
A. He doesn't provide reason B. His reason is: (please mark using dotted line)
2. Bob's evidence is fact d, which of the following is true of his evidence? ( ) (one or more choice)
A. Bob's evidence is relevant to his claim B. Bob's evidence is sufficient to prove that his claim is right C. None of the above
3. Jane's rebuttal against Bob is marked using the straight line, which of the following is true of her rebuttal? ( ) (one or more choice)
A. Jane points out Bob's deficiency B. Jane proves Bob's deficiency using appropriate evidence C. Jane provides her own claim and explains it D. All of what Jane says are right E. None of the above

Figure 4. An example of a task before and after modification

DISCUSSION

In this study, we presented influencing factors found in an iterative instrument development process that contribute to the improvement of SAC assessment. Based on the findings of the study, we conclude that the iterative process is reasonable in designing the instrument, both teachers and students provided useful feedback. Besides, the instrument performed better in the field after considering these factors and brought positive consequences to participants. Students were interested in taking this "unusual formed" test and felt much less confusing and provided more positive feedback about interacting with the modified instrument that was revised according to the factors shown above. Although students seldom been engaged in SA, they found it helpful in promoting their thinking ability and learning. We argue that this study provides detailed information in supporting SAC assessment research, and further studies are encouraged to test its application in other contexts.

REFERENCES

Diana, N. Y. P. (2019). Assessment of Argumentation in Chemistry: A Model for Designing Items. In *Argumentation in Chemistry Education* (pp. 106-141).

Newton, P. E. (2017). An approach to understanding validation arguments. *Ofqual, Coventry*.

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science education*, *88*(6), 915-933.

Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, *94*(5), 765-793.

Wilson, M. (2004). *Constructing measures: An item response modeling approach*: Routledge.

Toulmin, S. (1958). The uses of argument. Cambridge: Cambridge University Press.

Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentive competence. *Cognition and Instruction*, *31*(4), 456-496.

Osborne, J. F., Henderson, J. B., MacPherson, A., Szu, E., Wild, A., & Yao, S. Y. (2016). The development and validation of a learning progression for argumentation in science. Journal of Research in Science Teaching, 53(6), 821-846.